FSB FINANCIAL STABILITY BOARD

Sound Practices for Responsible Adoption of Artificial Intelligence (AI): Consultation report

Response to Consultation

Fintech Open Source Foundation (FINOS)

1. **Do you agree with the benefits and risks of AI adoption by financial institutions described in this report? Are there any substantive benefits or risks not covered?**

FINOS agrees with the report's identification of third-party dependency, service-provider concentration, and cyber vulnerabilities. AI compresses the timeline from vulnerability discovery to exploitation.

Furthermore, agentic AI introduces distinct risks; agent autonomy can create or amplify risks at great speed, including unauthorized actions, erroneous execution, and memory poisoning, alongside the challenge of overseeing agents acting as "synthetic employees".

Additionally, the final report should emphasize risks arising from the architecture of agentic systems—not merely from the underlying model, as well as risks from architectural drift.

To directly address these vectors, the open-source FINOS AI Governance Framework (AIGF) (https://air-governance-framework.finos.org) details specific agentic risk profiles:

RI-24: Agent Action Authorization Bypass.

RI-25: Tool Chain Manipulation and Injection.

RI-26: MCP Server Supply Chain Compromise.

RI-27: Agent State Persistence Poisoning.

RI-28 & RI-29: Multi-Agent Trust Boundary Violations and Credential Harvesting.

In enterprise architecture, deployed systems frequently diverge from approved designs—a risk amplified when coding agents generate code. Members of FINOS address this using CALM (Common Architecture Language Model) for expressing Architecture-as-Code (AasC). CALM (calm.finos.org) plays a vital role in Governance, Risk, & Compliance (GRC) processes at banks by 1) allowing development teams to receive approval for an architecture pattern, which allows for significantly reduced time for approval to deploy (https://www.ciodive.com/news/morgan-stanley-open-source-software-development-finos-calm/758339/), and 2) CALM's JSON-spec schema allows for the embedding of Controls at

the component level of architecture diagrams for threat modeling & evidencing (https://www.infoq.com/news/2026/03/morgan-stanley-apis-mcp-calm/).

Utilizing AasC allows firms to ensure that the software architecture that is deployed matches the architecture that was approved by GRC. CALM utilizes strict versioning to maintain an auditable history of approved architecture patterns, allowing institutions to continuously validate deployed code against specific approved versions and resolve the risk of GRC-to-implementation disconnect.

The FINOS AI Reference Architecture working group (https://github.com/finos/ai-reference-architecture-library) has also codified AI Architectures associated with FSI use cases in CALM, and then modeled the architectures with Risks & Mitigations from the AI Governance Framework. FSI Vendors (e.g. https://docs.moderne.io/user-documentation/agent-tools/prethink/) utilize CALM to ensure that AI generated code matches approved architecture patterns.

2. **Are the sound practices sufficiently comprehensive and clear to enable financial institutions' responsible AI adoption?**

While high-level principles are sound, translating expectations into bespoke controls creates duplicated costs and non-comparable assurance. The FSB should recognize open, machine-readable control libraries ("governance as code") as a legitimate means to operationalize the sound practices.

FINOS provides several open-source implementations:

AI Governance Framework (AIGF) [https://air-governance-framework.finos.org/]: Pairs an AI risk catalogue with deployable mitigations.

Common Cloud Controls (CCC) [https://finos.org/common-cloud-controls]: Establishes consistent controls across major cloud providers to address SP 11 and 12 concentration risks.

OSERA [https://osera.finos.org/]: Mutualizes time-bound open-source component backpatching under vendor-neutral SLAs to mitigate supply-chain vulnerabilities.

Common Domain Model (CDM) [https://www.finos.org/common-domain-model]: Standardizes financial product representations to improve data quality, lineage, and provenance under SP 7.

Specific Sound Practice Amendments based on AIGF mitigation catalogue:

Decision Traceability vs. Chain-of-Thought (SP 8 & 9): Replace "chain-of-thought logging" with decision and action traceability, as stated reasoning may not reflect actual model execution. This aligns with AIGF mitigations MI-21 (Agent Decision Audit and Explainability) [https://air-governance-framework.finos.org/mitigations/mi-21_agent-decision-audit-and-explainability.html], MI-13 (Source Traceability) [https://air-governance-framework.finos.org/mitigations/mi-13_providing-citations-and-source-traceability-for-ai-generated-information.html], and MI-4 (System Observability) [https://air-governance-framework.finos.org/mitigations/mi-4_ai-system-observability.html].

Automated Safeguards (SP 11 & 12): Operationalize regulatory policies using machine-readable controls like AIGF MI-18 (Least Privilege Framework) [https://air-governance-framework.finos.org/mitigations/mi-18_agent-authority-least-privilege-framework.html] and MI-20 (MCP Server Security Governance) [https://air-governance-framework.finos.org/mitigations/mi-20_mcp-server-security-governance.html].

3. **Do the sound practices strike an appropriate balance between managing risks relating to all forms of AI, and addressing some of the risks relating to emerging and new complex forms of AI, such as GenAI and agentic AI?**

   Agentic AI is the urgent frontier requiring open, shared scoping conventions—specifically autonomy levels, action classes, and control-environment descriptors—so assurance evidence remains comparable across firms and supervisors. Assurance must be scoped to defined contexts rather than unbounded "safety" claims.

   Baseline agentic controls should include machine identities, tool allowlisting (https://csrc.nist.gov/glossary/term/application_allowlisting) , human approval checkpoints, memory integrity controls, and auditable tool access logs. FINOS supports this via the FINOS AI Fund (collective investment in agentic governance) and open orchestration tooling via Fluxnova (fluxnova.finos.org) that provides concrete oversight primitives (approval checkpoints, dual authorization, kill-switches, audit trails). AIGF MI-19 (Tool Chain Validation and Sanitation) (https://air-governance-framework.finos.org/mitigations/mi-19_tool-chain-validation-and-sanitization.html) provides best practices on tool chain validation and execution sequence sanitization.

4. **Are the sound practices sufficiently flexible to accommodate and address newer types of AI and responsible adoption over time?**

   To remain flexible, the FSB should encourage institutions to adopt versioned, living risk taxonomies maintained by neutral open-source consortia. Applying strict versioning to resources like the AIGF ensures financial institutions can systematically update controls while preserving an auditable history of their governance posture over time.

   Furthermore, FINOS is building an open-source pipeline using Gemara, Fluxnova, OpenTelemetry, Grafana, CALM, and CCC to supply the technical infrastructure needed for real-time observability and continuous control feedback loops.

5. **Do the case studies in this report sufficiently highlight how different types of financial institutions can benefit from responsible AI adoption? Are there additional case studies for inclusion in the report? If so, please provide such case studies, particularly for nonbanks.**

   The FSB should engage neutral open-source consortia to develop additional case studies for nonbanks and financial market infrastructures.

   Additionally, FINOS offers the Governance-as-Code Loan Approval Demo (https://www.finos.org/hubfs/OSFF%202026%20(Open%20Source%20in%20Finance%20Forum)/OSFF%202026%20London/Videos/AI%20Governance%20Unlocked_%20The%20Steel%20Thread%20Demo%20-

%20James%20McLeod%2c%20NatWest%20Group%20%26%20Olivier%20Poupeney.mp4) as an actionable example. It demonstrates end-to-end GRC execution by combining AI risk controls (AIGF), Architecture-as-Code (CALM), and enterprise orchestration (Fluxnova) to yield an auditable AI workflow (https://ai.finos.org/governance-as-code)

6. **Do the case studies in this report provide actionable insights for financial institutions in their responsible AI adoption?**

   No comment

7. **Are the definitions in the glossary clear and aligned with industry sound practices, including recent developments in AI?**

   FINOS recommends the following glossary refinements:

   Agentic AI: Define explicitly by planning, tool use, execution, persistence, and interaction with external systems, rather than defining it primarily by limited human oversight.

   Open-Source vs. Open-Weight: Explicitly distinguish open-source AI from open-weight models; releasing weights alone does not provide access to training code, data, or open-source software rights.

   Term Distinctions: Maintain clear distinctions between explainability, transparency, monitoring, observability, and auditability.

   Recommended Additions: Add definitions for AI agent, multi-agent system, tool interface, orchestration, persistent memory/state, human-in/on-the-loop, Model Context Protocol (MCP), AI benchmarking, and AI red-teaming.

8. **Are there any comments you would like to make on this report? Please fill in.**

   A letter from FINOS was sent as supplementary material.

   Here is a list of sources & further reading:

   A letter from FINOS was sent as supplementary material.

   FINOS AI Governance Framework (AIGF)

   https://ai.finos.org/

   FINOS Governance-as-Code / Loan Approval Demo

   https://ai.finos.org/governance-as-code

   AIGF Mitigation MI-4: AI System Observability

   https://air-governance-framework.finos.org/mitigations/mi-4_ai-system-observability.html

AIGF Mitigation MI-13: Providing Citations and Source Traceability

https://air-governance-framework.finos.org/mitigations/mi-13_providing-citations-and-source-traceability-for-ai-generated-information.html

AIGF Mitigation MI-21: Agent Decision Audit and Explainability

https://air-governance-framework.finos.org/mitigations/mi-21_agent-decision-audit-and-explainability.html

FINOS CALM (Common Architecture Language Model)

https://calm.finos.org/

FINOS Fluxnova

https://fluxnova.finos.org/

FINOS Common Cloud Controls (CCC)

https://finos.org/common-cloud-controls

FINOS Common Domain Model (CDM)

https://finos.org/common-domain-model